Scientific Journal of Impact Factor (SJIF): 4.72

e-ISSN (O): 2348-4470
p-ISSN (P): 2348-6406

International Journal of Advance Engineering and Research Development

Volume 4, Issue 12, December -2017

# A STUDY ON DNA ENCRYPTION

K.SRIVIDHYA[1] & M.DEVI[2]

[1]Scholar,BTech ECE,SRC,Sastra University,Thanjavur, India
[2]Assistant Professor,SRC,Sastra University,Thanjavur,India

**Abstract.** *With rapidly advancing technology and the prospect of all data transmission going digital soon, secure data transmission is indeed a subject to be broached over.DNA encryption provides a valuable hand in providing information security and addresses various aspects like data confidentiality, data integrity and authentication. Various techniques of DNA encryption have now come to practice like Polymerase Chain Reaction, Electrophoresis and Stegonography.*

## 1. INTRODUCTION

Cryptography is a technique of using secret codes for secure data transmission in the presence of third parties. It involves creation of encryption algorithms that scramble data and convert it into cipher text that is undecipherable by third parties. At the receiver end, a key is used to convert the cipher text into original data using a decryption algorithm. Cryptography provides integrity, authentication and confidentiality. Integrity involves protection of data from impersonation using a technique involving digital signatures. Authentication provides access to digital ID'S of parties to verify if they are authentic or not. Confidentiality keeps the information secret by converting it into a cipher text.

## 2. TECHNOLOGIES USED IN CRYPTOGRAPHY

### 2.1 PCR

In this technique, a single segment of DNA is amplified to generate millions of replicas of the DNA sequence. It is used to detect the presence of pathogens in a host organism by amplification of nucleic acids. Two sets of primers and the enzyme DNA polymerase are used to generate millions of copies of a gene. The amplification of the DNA segments is facilitated by Taq polymerase which is active when denaturation of DNA occurs at high temperatures.

### 2.2 DNA CHIP TECHNOLOGY

In this technique, segments of DNA called DNA spots get attached to a solid surface. This technique is used to hybridise CDNA or CRNA.DNA chip technology is mainly used in expression profiling where expression levels of genes are studied. They are used in the treatment of certain diseases.

### 2.3 GEL ELECTROPHORESIS

The segments of DNA are fragmented using restriction endonucleases.The separation of DNA fragments is done using a matrix like agarose gel.The DNA fragments get separated according to their size and the separated fragments are then removed from the agarose gel.These separated DNA fragments can be purified and used to generate recombinant DNA.

### 2.4 DNA FRAGMENT ASSEMBLY

This technique involves merging fragments of DNA to get back the original DNA sequence.Using shotgun sequencing,the DNA is broken into small segments and by continuous fragmentation and sequencing,many overlapping reads of DNA are obtained which are then reassembled to get back the original DNA sequence.

## 3. A GLIMPSE OF DNA ENCRYPTION

DNA encryption involves secret transmission of data using the four nitrogenous bases in a DNA where each base corresponds to a two digit binary value. Symmetric encryption is used for secure data transmission in DNA cryptography. Unlike asymmetric encryption where a key is used to encrypt the message at the sender end and another key is used to decrypt the message at the receiver end, symmetric encryption uses private and public keys. Both sender and receiver use the same secret key/public key to encrypt and decrypt messages. Apart from this, the sender also shares a common key/private key with the receiver to recover the plain text back from the cipher text.

 

## 4. THE DNA DOUBLE HELIX

DNA is a nucleic acid that acts as the genetic material in most of the living organisms. It contains a chain of deoxyribonucleotides. A nucleotide is composed of nitrogenous base, a pentose sugar and a phosphate group.The nitrogenous bases are of two kinds: purines(Adenine and Guanine) and pyramidines(Cytosine,Thymine and Uracil).Thymine is a component of DNA while Uracil is a component of RNA.The backbone in a nucleotide chain is formed by sugar and phosphates with nitrogenous bases projecting from the backbone.

James Watson and Francis Crick proposed the double helix structure of DNA that explained the pairing of Adenine and Guanine and Thymine and Cytosine in the polynucleotide chain. This base pairing of polynucleotide chains indicates that they are complementary to each other. Adenine forms a double bond with Thymine while Guanine forms a triple bond with Cytosine. The base pairing of polynucleotides put forth by double helix model found support from the observations of Erwin Chargaff that suggested that the ratios between A and T and C and G are constant and equal to one.

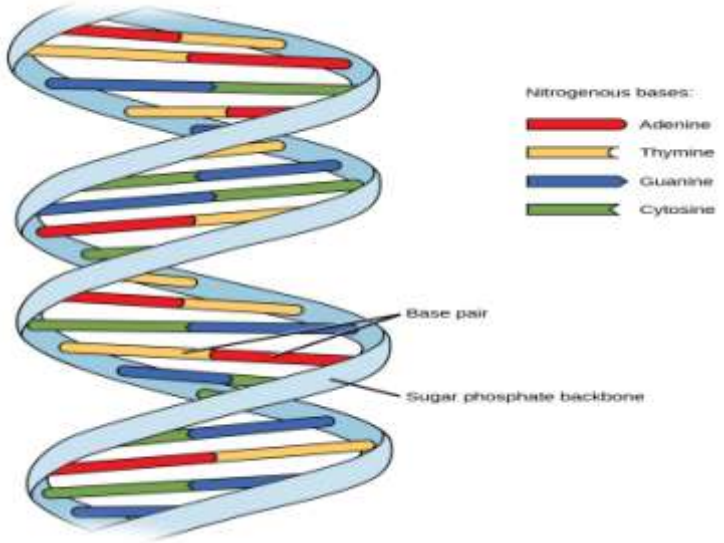


Figure 1: Double helix model of DNA

5. ENCRYPTION BASED ON DNA SEQUENCES

The central dogma of molecular biology states that flow of genetic information takes place from DNA to RNA to proteins.

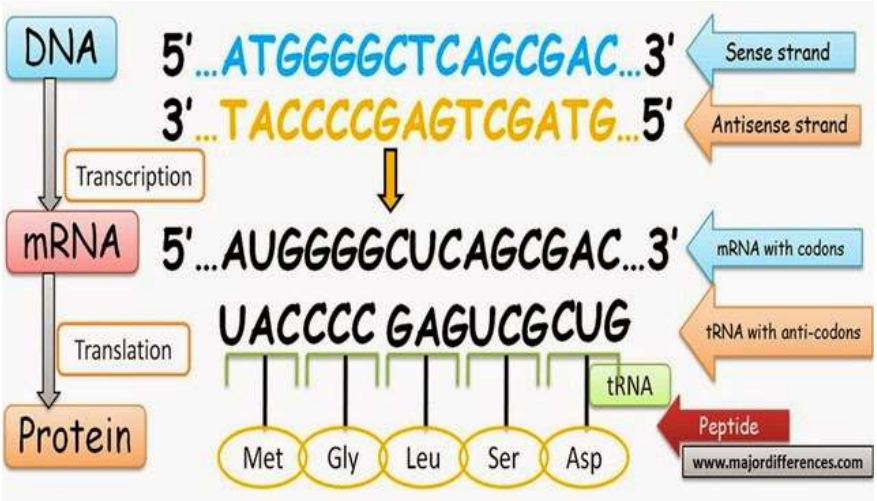


Figure 2: Central dogma of molecular biology

There are three important processes in the conversion of a DNA sequence into protein: replication, transcription and translation.

.The replication of DNA is facilitated by an enzyme called DNA dependent DNA polymerase. This is followed by transcription that involves copying of genetic information from one strand of DNA to another. The two strands of DNA have

 

opposite polarity. The DNA strand that has a polarity 3'→5' acts as the template and is called template strand/sense strand. The other strand of DNA with 5'→3' polarity is called coding strand/anti sense strand and is displaced during transcription. The coding sequences in a DNA double helix are called exons. The exons are interrupted by introns that get removed in the processed RNA.mRNA and tRNA are also vital in the process of transcription. mRNA provides the template while tRNA is used to represent the genetic code in terms of amino acids. The process of transcription involves the formation of mRNA sequences from the template strand of DNA and splicing of the coding strand of DNA. The process of transcription is followed by translation where polymerization of amino acids occur to form a polypeptide. The genetic sequence is expressed in groups of three nucleotides called codons where each codon codes for one amino acid.

## 6. PROPOSED ALGORITHM

A DNA computing based encryption and decryption algorithm has been used.
The four bases in a DNA sequence are represented in digital form for encryption as represented below.

| NUCLEOTIDE | BINARY FORM |
|---|---|
| A | 00 |
| T | 01 |
| C | 10 |
| G | 11 |

Table 1: Binary representation of nucleotides

6.1 STEPS INVOLVED IN ENCRYPTION

The original DNA sequence is first encoded at the sender end. It can be represented as
3'→5' CTAGCTAGCTAG (template)
5'→3' GATCGATCGATC (coding)
This is followed by transcription of message sequence into mRNA. The introns are removed and only exons participate in transcription.
CTAGCTAGCTAG →CUAGCUAGCUAG (transcription)
Now, the mRNA gets translated into protein. The resulting sequence can be represented as a chain of amino acids where each amino acid is formed by a combination of three bases called codons.
CUAGCUAGCUAG→GAUCGAUCGAUC (translation)
(Asp Arg Ser Ile)
The resulting sequence of amino acids acts as a public key to the receiver. The sender also sends a private key providing information about the location of introns. The formation of amino acids from bases is depicted in the table below.

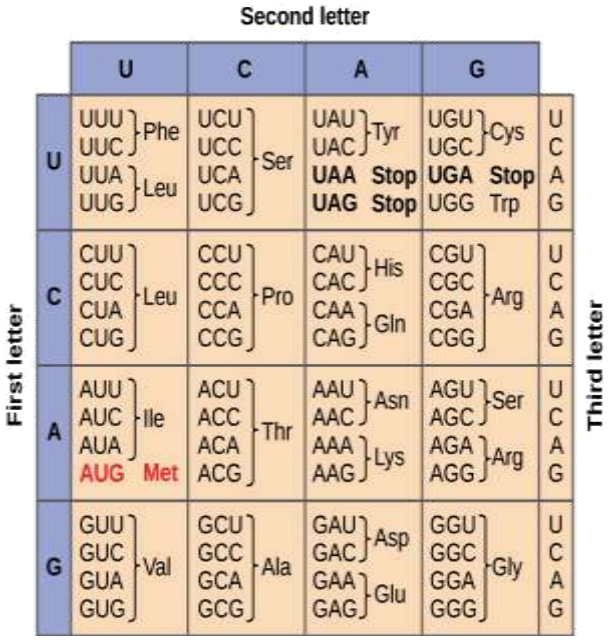


Figure 3: Codons for amino acid

6.2 STEPS INVOLVED IN DECRYPTION

The receiver receives proteins from the sender and uses the public key to decrypt the original DNA sequence. The tRNA sequence GAUCGAUCGAUC can be obtained from the sequence of amino acids through reverse translation. This is followed by the recovery of mRNA sequence CUAGCUAGCUAG through reverse transcription.The template strand of DNA sequence CTAGCTAGCTAG can be obtained from the mRNA sequence. The location of introns can be detected by the private key provided by the sender and the original DNA sequence can hence be decrypted.

## 7. CONCLUSION

In this paper,DNA cryptography has been explained using DNA computing based encryption and decryption algorithm. Symmetric encryption using private and public keys has been used in this technique to transfer data confidentially in the presence of third parties.

## 8. REFERENCES

[1]Anupriya Aggarwal,Praveen Kanth,"Secure data transmission using DNA encryption",Computer Engineering and Intelligent Systems,Vol.5,No.7,2014

[2]Grasha Jacob,A.Murugan,"DNA based Cryptography:An Overview and Analysis",Int.J.Emerg.Sci.,3(1),36-42,March 2013

[3]Martyn Amos,Gheorghepaun,Grzegorz Rozenberg,Arto Salomaa,"Topics in the theory of DNA computing",Theoritical Computer Science 287(2002)3-38

 